DOCUMENT RESUME

ED 179 385 SE 029 194

AUTHOR Page, R. L.; And Others
TITLE Attitude Assessment in Science and Technology. Occasional Paper.
INSTITUTION Bath Univ., (England). Science and Technology Education Centre.
PUB DATE 79
NOTE 50p.; Not available in hard copy due to copyright restrictions; Contains marginal legibility in Appendix 2
AVAILABLE FROM Science & Technology Education Centre, University of Bath, Claverton Down, United Kingdom BA2 7AY (3.50 pounds)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Affective Tests; Attitudes; *Attitude Tests; Industry; *Instrumentation; Measurement Techniques; Research; Science Education; Sciences; *Science Tests; Secondary Education; Technology; *Test Construction

ABSTRACT

This paper, intended as an introduction to attitude scale construction, outlines the theory behind attitude measurement and describes the Thurstone and Likert methods of attitude scale construction. The methodology is illustrated with examples from the construction of a scale the authors developed to measure attitudes to technology and industry. References to texts which supply more detailed information are included. The authors' scale and a description of its construction are included in the appendix. (BT)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
Duncan Harris
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED179785

ATTITUDE ASSESSMENT in Science and Technology

R.L. PAGE M. MASTERTON

Occasional Paper

SE 029 194

ATTITUDE ASSESSMENT in Science and Technology

R.L.Page, M. Nash, M.Orr

Occasional paper
published by

Science and Technology Education Centre
University of Bath
Claverton Down
Bath
BA2 7AY
UK

Published and distributed by the

Science and Technology Education Centre
University of Bath
Claverton Down
Bath
BA2 7AY

©

R.L. Page, M. Nash, M. Orr 1979

Other than permitted under the Copyright Act 1956, no part of this publication may be photocopied, recorded, or otherwise reproduced, stored in a retrieval system or transmitted in any form by any electronic or mechanical means without the prior permission of the copyright owner.

First published 1979

ISBN 086197 000 4

ATTITUDE ASSESSMENT IN SCIENCE AND TECHNOLOGY

Introduction

In the educational field an increased interest in measuring the attitudes of school pupils in relation to science and technology has emerged. The assessment of scientific and technological potential, once restricted to the measurement of cognitive and psychomotor skills, has widened to include attitude assessment, with the awareness of the effect attitudes have upon cognitive and psychomotor abilities. The construction of attitude scales to measure pupils' attitudes to scientific and technological subjects has been one natural outcome. Such scales have wide educational application: they can be used to evaluate the effect of a new course, new equipment or new teaching techniques on the attitudes of pupils studying science and technology. They can also be used to assess differences in attitudes between pupils with differential science and technological experience. They have their limitations as well, however, and these will be expanded on in the text. Most importantly they should not be used to assess an individual's attitude, neither should decisions about an individual, whether vocational or otherwise, be made on the basis of a pupil's "score" on such a scale. Bearing in mind these limitations, this paper aims to give a brief outline of the theory behind attitude measurement and an equally brief description of the two most commonly used methods of attitude scale construction, and modifications to these methods. This will be illustrated where appropriate, with examples from the construction of a scale the authors developed to measure attitudes to technology and industry. The intention is to provide a useful introduction to the subject for anyone contemplating constructing a scale themselves, and for the general reader, a basic understanding of attitude scales. References to texts which supply more detailed information are included, for the benefit of those who wish to pursue the subject in greater depth. The author's scale and a description of its construction, is included in the appendix as an example of the end product of some of the techniques described in the paper.

CONTENTS

TEXT

		Page No.
1.	What is an attitude ?	2
2.	How do we measure attitudes ?	3
3.	The Construction of attitude scales	7
4.	The construction of Thurstone scales	13
5.	Further attempts to ensure unidimensionality	17
6.	The final Thurstone scale	17
7.	The construction of Likert scales	19
8.	The final Likert scale	21
9.	Comparison of ease with which Likert and Thurstone scales are constructed.	23

APPENDIX

1.	The development of the technology scale.	26
2.	Attitude to technology and industry scale.	38

REFERENCES 41

1. WHAT IS AN ATTITUDE?

A. N. Oppenheim's [12] definition of an attitude seems to be representative of the many definitions available in texts on the subject. It is a "tendency to act or react in a certain manner when confronted with certain stimuli". The stimuli may be an object, institution, person, idea, ideal or anything else that a person can react negatively or positively towards. The reaction may also take many forms - physical, mental, verbal and so on.

It is generally believed that a person's attitude towards an object lies dormant most of the time, and becomes expressed in speech or behaviour, only when the object of the attitude is perceived, either by being seen or alluded to in some way.

For example, a person who has an unfavourable attitude towards the use of atomic energy, who hears the news that another nuclear power station is to be set up will express his attitude at least by mental disapproval, and probably in a more physical or verbal manner.

Isolating attitudes from other personality variables

In order to illustrate more clearly what an attitude is, psychologists often separate conceptually attitudes from other closely related components of behaviour. These are then isolated and described in turn to show what an attitude is not. Whilst a brief description of these other personality variables will be given here, it must be born in mind that this is an artificial separation and fails to do justice to the unique combinations of attitudes each person has, and their relationship to other personality variables.

Attitudes, then, are thought to be separate from, but may be connected to, a person's underlying values and indeed, their whole philosophy of life. Attitudes may produce motives, which tend to be expressed by goal-directed behaviour to achieve or do various things. Attitudes may be re-inforced by beliefs, which are verbalized in the form of opinions. Beliefs are said to be different from attitudes because beliefs are held cognitively, as the result of knowledge, and without emotional commitment. Attitudes, in comparison, are affective, and the attitude response to stimuli is an emotional response. Whilst a person's beliefs might be changed by the evaluation of new evidence which has been presented to him, a person's

attitude will only be changed if he comes to accept that the attitude object has negatively or positively valued characteristics which the person had not previously accepted.

It is thought that attitudes are learned by conditioning, suggestion and imitation, from family and home, and to a lesser extent, peer group and educational or vocational establishment.

2. HOW DO WE MEASURE ATTITUDES?

The actual response to stimuli, produced by an attitude, is not recognised as part of the attitude itself. However, the only way we can gain knowledge of a person's attitude is to examine his response.

For the purposes of social research it is often necessary to try and gauge a person's attitude, and attitude scales are frequently used to do this. The reasons why alternative methods are thought unacceptable will be outlined here and then the arguments advocating an attitude scale as the best instrument to measure attitudes will be put forward.

If we want to know a person's attitude to an issue we could just ask him how he feels about it. This has certain disadvantages; if we have asked him for his attitudes towards a controversial issue he may not voice his true feelings but give what he considers to be a socially acceptable reply. Also, he may be aware of our attitude to the subject, or may be able to hazard a guess as to our attitude from our ages, appearance, accent or social position, and so give a response he considers will please us. Another objection to this method is that a person may not be aware of his true feelings towards an object or his feelings towards an object may be ambivalent. Most people seem able to tolerate considerable inconsistency in their feelings towards objects, issues, persons etc. and so it is unlikely that an individual will be able to weigh and evaluate the intensity of two opposed feelings and decide which is dominant or stronger. (However, it could be argued that an attitude scale does not adequately cope with this possibility either). A question may not, in any case, provide the right stimuli to evoke an attitude response, it is argued.

An alternative method of assessing a person's attitude would be to study their behaviour, since attitudes are frequently manifested by behavioural responses to stimuli. Unfortunately, from the point of view of the research worker this method has limited application, since social behaviour is often designed to conceal true feelings. Overt behaviour can be a poor guide to attitudes, as well as being very time consuming for the social researcher, who may have to wait a long time for the right stimuli to be presented to the right person in order to observe his behavioural response.

A further limitation of either of these methods, it is argued, is that whilst they may indicate whether a person has a favourable or unfavourable attitude to an object, they will not grade the degree of favourability or unfavourability with which the attitude is held.

This, it is hoped, is achieved by an attitude scale. An attitude scale is a series of statements relating to a clearly defined topic, expressing favourable, unfavourable, or neutral attitudes to it, and these statements are thought to provide the appropriate stimuli to evoke a reaction which is presumed to be produced by the attitude. The reaction is in the form of whether the person agrees or disagrees with the statements, or whether he neither agrees nor disagrees, and the extent to which he agrees or disagrees with the statements. His reactions are scored and the score gives the person's attitude as a position on a scale from favourable, through neutral to unfavourable. It should be noted that this process does constrict the person's reactions to stimuli - the exact statement a person would endorse may not be present in the scale, and there is no room for a physical response on the part of the respondent other than refusing to answer the scale and walking out of the room. Also attitude scales do not prevent respondents from attempting to portray themselves as possessing socially desirable attitudes.

It is generally accepted that attitude scales have other limitations as measuring devices. Criticisms tend to relate to the inaccessibility of attitudes to measurement however, rather than to deficiences in attitude scales themselves.

Attitude scales require the assumption that attitudes are unidimensional i.e. that every attitude can be isolated from every other attitude, and from other components of behaviour, such as beliefs, motives etc. It is arguable whether this is in fact so - a person's unfavourable attitude towards vivisection may be part of a wider unfavourable attitude towards the use of dead animals in scienctific experiments, or bound up with religious beliefs,etc. It is also doubtful whether attitude scales are unidimensional. Various procedures are adopted in the construction of an attiude scale (described later) to attain this goal, but it is unlikely that it is ever achieved. Shaw and Wright [16] report moderately high correlations between scales purporting to measure different things, and less than perfect correlations between scales designed to measure the same thing. Some authors have pointed out that an absolutely unidimensional scale would probably have to consist of statements which were almost identical in meaning and hence repetitious - in which case a unidimensional scale could only realistically consist of one statement. For the rest of this paper, however, unidimensionality will be taken to mean closely intercorrelated items with homogeneous content.

Scores on attitude scales are not usually reproducible - partly because they are not unidimensional. In other words, a person's score on an attitude scale can often have been arrived at in a number of different ways, and does not tell us which statements the respondent has agreed with and which he has not. So it is impossible to plot the exact progression of an attitude from unfavourability to favourability, or vice-versa. Some methods have been devised to overcome this problem, e.g. the Guttman technique (Guttman scales tend to consist of statements virtually interchangeable in meaning) but always at the expense of some other requirements of an adequate measuring instrument. As a result, it is generally necessary to look at the pattern of a respondent's answers, as well as his overall score.

Attitude scales also require the assumption that attitudes range along linear continuums. This probably imposes an artificial model on attitudes: the favourable part and unfavourable part of an attitude may not be linear extensions of one another, a favourable attitude need not be the opposite of an unfavourable attitude.

To use an example of Oppenheim's [12] from a scale to measure the attitude of mothers to their children, rejection of a child may not be the exact opposite of acceptance, since other components may enter into the attitude of the rejecting mother. It would be possible for a mother to have an ambivalent attitude to her child, in which case she should score high on both the acceptance and rejection ends of the scale. So graphic representations of attitudes might be expressed more accurately as concentric circles, overlapping elipses or curvilinear lines. However, since there is no evidence to suggest that a model other than linear is a more accurate representation, and a linear model is most useful for measurement, those constructing an attitude scale must attempt to measure the negativity-positivity of a single attitude, along a straight line.

Attitudes are probably not, in many cases evenly distributed along a continuum. The degree of differentiation at one end of an attitude continuum may be very different from that at the other end; for instance Seashore and Hevner point out that attitudes towards music may be mostly distributed along a range from enthusiasm to indifference, with few people having a strongly unfavourable attitude. In consequence, attitude scales cannot assume equality of units, and are almost always on ordinal scales. In other words, the differences between a score of 20 and 30 is not necessarily the same as the difference between a score of 50 and 60.

Consequently it is not possible to compare the scores of different individuals on the same scale, or to average a person's score on a number of scales. The best that can be done is to rank people according to the degree of favourability or unfavourability of their attitude. However, the differences between the average scores of separate groups can be compared and changes in the attitude of a group as a function of experimental treatments can be assessed.

Another problem connected with the measurement of attitudes is that attitudes can be held with greater or lesser intensity. Attitudes at the extremes of the attitude continuum are generally held with the greatest intensity, whilst the more neutral position is defended with far less vehemence. So there is a U-shaped

relationship between the attributes of intensity and content. This has led to the suggestion that to find the neutral point on a scale one must look to the point of minimum strength or intensity. (Lack of intensity of feeling, however, does not guarantee lack of negative or positive attitude). The identification of a neutral point on an attitude continuum is, in any case, rather arbitrary. A person labelled as having a neutral attitude may in fact have a slightly negative or positive attitude, or neutrality may still predispose or deter a person from verbal or physical behaviour. It has been argued, in fact, that attitudes must always be negative or positive; that if a person is completely neutral towards a stimuli then he has no attitude at all.

Methods of assessing the reliability or validity of attitude scales will not be dealt with in this paper, since the procedures are basically those employed in evaluating any psychological measuring instrument. However, it is worth mentioning that whilst most attitude scales are found to be sufficiently reliable, with reliability coefficients of .80 or higher, there are, as with all psychological tests, no truly adequate methods for assessing their validity, and indeed Shaw and Wright [16] would contend that the majority of attitude scales constructed have been insufficiently validated.

3. THE CONSTRUCTION OF ATTITUDE SCALES

There are many methods for constructing attitude scales. However, the majority of attitude scales have been constructed either by a method developed by Thurstone & Chave in 1929 - the 'method of equal-appearing intervals', usually just called the 'Thurstone method', or by the method developed by Likert in 1932, the 'method of summated ratings', usually called the 'Likert method' or slight modifications of these. More recent techniques of construction appear more complicated, and because fewer people have used them it is difficult to ascertain if these later methods offer any real improvements over the earlier methods.

The description of how to construct attitude scales in this paper, is confined to the Thurstone and Likert methods, together with a few refinements and short cuts developed for them, although

many of the points mentioned would be applicable to other techniques of construction. Where capital letters appear in the text this refers the reader to sections in the appendix so lettered, for an example of where the process described in the text was used in the construction of the authors' scale.

An attitude scale may be constructed on any subject which it is possible to have a negative or positive affective reaction towards.

Having decided to construct an attitude scale on a certain topic, in the authors case towards technology and industry, the initial procedures for the construction of Thurstone or Likert scales are very similar. Firstly, it is necessary to:

Explore the Origin and Extent of the Attitude

In other words, we need to get some idea of how the attitude arose initially, from whence it was acquired, how it intermingles with other attitudes and personality variables, whether it is specific or wide-ranging, whether it is split-up into sub-areas and so on. To use an example the authors are familiar with, i.e. attitudes of children towards technology, one needs to know whether attitude towards technology arises from experience of technological subjects at school, or the influence of T.V., or parents and teachers views, or whether it is part of a general attitude towards industry, or an appreciation of the benefits or dangers of the products of technology.

To do this it is advisable to read literature relevant to the subject of the attitude scale, ask experts concerned with the subject of the attitude scale for their opinions on the subject (e.g. in our example technologists, teachers and instructors etc) and most importantly, conduct in depth unstructured interviews with samples taken from the population the final scale is intended for, or ask them to write about how they feel about the subject. Such interviews and essays should be unstructured, so that feelings are not "suggested" to the respondents, but volunteered by them.

From the interviews, etc. we should have a better idea of what function the attitude scale is to play in our research, and

from this, what exactly we want the attitude scale to measure.(A) If we keep in mind the purpose of the scale, we shall avoid the mistake Gardner [6] notices in many attitude scales, of items being included which bear no relation to what it is the scale constructors wish to measure.

We should also have a better idea of how the attitude is distributed over the attitude continuum, and if there appears to be several factors involved in the attitude towards the subject under scrutiny. To use our previous example, having set out to construct an attitude scale towards technology, if it is found that this attitude splits up into attitude towards industry, attitude towards technological subjects, attitude towards a technical training, and so on, then it is safer to treat these aspects as separate attitudes, to be measured on separate sub-scales, unless, and until, it is possible to show that they are too highly correlated to be meaningfully separated.(B)This may be revealed by factor analysis, or by correlational or covariance statistics, but it is unlikely to be conclusively established at this stage in construction (methods of establishing the homogeneity of scales will be discussed later). The sub-scales then, will be treated as parts of an overall scale to measure attitude towards technology. Such sub-scales would have to be scored separately, and the total scores for each sub-scale should not be added together to give an overall score, as this score might well be meaningless. The sub-scale total scores must instead be treated as measures of unique attitudes which are interlinked. Bearing this in mind, it is at this point in the construction of a scale that we should decide whether it is necessary to measure all these sub-areas, or whether to be selective, and this will again depend on the purpose we have in mind for the scale. In general, the more aspects of a particular attitude that can be included, the more likely we are to get scores that will mean something in terms of the underlying attitude.

Some scale constructors in the past have failed to establish whether their scales include factors requiring separate scales, and may have consequently included several factors within one scale.

Such scales cannot possibly be unidimensional, and if such scales are used to measure change in attitude as the result of

experimental treatment, they may well fail to record any changes, since change in one factor as part of a scale, may not change an overall score significantly, or may cancel out with a change in another factor in the opposite direction.

From a decision on what part the attitude scale is to play in our research should come a decision on what type of attitude scaling technique to utilize to construct the scale. However, as there are other preliminary procedures which Likert and Thurstone scales have in common, and as describing the differences in construction between Thurstone and Likert scales will make obvious why different scales better fulfil different needs, how to decide which scale to use will be dealt with after covering these other points.

Having conducted interviews etc. it is necessary to:

Collect an item pool of statements

As already mentioned, attitude scales consist of a series of statements from half-a-dozen, to 2 dozen or more, with which the respondent is asked to agree or disagree. These statements sample the attitude range so that from a persons' disagreement or agreement with some or all of the statements, it is possible to place him approximately on the attitude continuum. In the initial construction of an attitude scale it is necessary to have a great many more statements than appear in the final scale because, as only a limited number of them can be used in the final scale, it is necessary to pick, through statistical procedures, the best statements available to ascertain a person's attitude.

So from the literature and interviews must be collected a list of statements concerning the object under consideration. (C) About 100 is the usual number. As the statements are to form the stimuli which evoke the attitude response, the statements have to be written in a form conducive to producing this reaction. A number of criteria have been formulated by various writers for writing attitude statements thought to facilitate an attitude response.

Amongst these are that statements should be representative of the almost unlimited number of statements that could have been

written on the attitude subject and there should be roughly equal numbers of items dealing with each main aspect whether contained in one scale or several sub-scales. Statements should be meaningful to the respondents, and for this reason they should be put into the language of the respondents. This is why statements written from taped interviews and essays are most useful. Also a respondent is more likely to make his feelings known on issues which are emotionally loaded or potentially embarrassing if they are mentioned in colloquial terms which put them at their ease. Attitudes concern emotions, therefore attitude statements should be emotional statements, so that they can differentiate between people with favourable and unfavourable feelings. Factual statements, whether they favour an object or not, are more likely to be endorsed by people with both favourable and unfavourable attitudes.

Edwards, [3] in his book 'Techniques of Attitude Scale Construction', supplies a list of "do's and don'ts" concerning attitude statements. The most important things are that statements are clear and unambiguous and that statements do not make respondents think too hard, as it is their immediate response which is required for measurement. (D)

To avoid 'Response Set', a tendency to reply to all items in a particular way - for instance to agree with all statements that appear to reflect a socially desirable attitude, positively and negatively worded items should be included in the scale, and the agree and disagree response categories should be reversed. Contradictory responses may then indicate a response set, so the respondents' replies can be viewed with caution.

Most textbooks suggest that statements collected for the item pool should express attitudes ranging from one extreme to the other of the attitude continuum. In the authors' opinion, however, the type of statements one collects should vary between the type of scale construction to be utilized.

A Thurstone scale is a short list of, say, 14 statements which range from very favourable toward the attitude object, through neutral, to very unfavourable. Each statement is scaled, so that each statement expresses a little less or a little more

favourability to the attitude object, than the statement preceding or following it. Each statement therefore, has a scale value i.e. where it is placed on the attitude continuum, and this scale value has been derived, and can be expressed, numerically by the author of the scale. When a person ticks agreement to a statement, he is in fact assigning himself a position, which can be numerically expressed, on the attitude continuum. Because of this, statements collected for Thurstone scales should sample the whole range of the attitude continuum.

A Likert scale, by comparison, is again a list of statements expressing attitudes towards an object, but in this case the respondent ticks one of a number of response categories, usually five, ranging from "strongly agree", "agree", "uncertain", to "disagree", and "strongly disagree". The number of response categories can of course be increased to 6, "strongly agree", "agree", "mildly agree", "mildly disagree", "disagree", "strongly disagree", in which case the respondent is 'forced' to have an attitude in one direction or the other, however mild, or the number of categories could be reduced to four or three. Respondents, therefore, place themselves on an attitude continuum for each statement, and a numerical scale position can be assigned from this. If we decide that a high score represents a favourable attitude, then if the item expresses a favourable attitude, "strongly agree" can be given a weight of 4, "agree" a weight of 3, "uncertain" 2, "disagree" 1, and "strongly disagree" 0. If the item expresses an unfavourable attitude the scoring system is reversed. Once a respondent has responded to all the statements in the scale, the response weights he has been assigned for each statement can be summed to obtain a total score, hence the name given to this method by Bird, the 'method of summated ratings'. If there were 15 statements in a scale and 5 response categories, a person could have a score of anything between 0 and 60. If a person had a very low score he would be considered as having a very unfavourable attitude towards the attitude object, and if he had a very high score, he would be considered as having a very favourable attitude. A middling score, of around 30 would be considered to represent a neutral attitude.

It will be obvious to the reader that the item pool of a Likert scale must consist of items which clearly express favour-

able, and unfavourable attitudes towards the attitude object, and not degrees of favourability, unfavourability and neutrality, or a nonsense will be made of the scoring procedure. For example there would be no appropriate score for a "strongly agree" response to a neutral item. However, if the item pool of a Likert scale contains neutral or near neutral items they would probably be eliminated at the item analysis stage (see later). Other problems connected with this scoring method will be discussed later in the paper.

4. The Construction of Thurstone Scales

Statements should be reproduced (a useful suggestion is to put them on cards, one card for each statement) and each statement assigned its own item number. The statements should then be given to a number of judges (Thurstone used over 300, but research has shown that 50 will do just as well). Probably any reasonably intelligent person would do as a judge, although Oppenheim [12] thinks that judges should be people similar to those for whom the final scale is intended. This may not be possible, of course, if the scale is intended for children.

Each judge should be given a complete set of statements in random order, told the subject matter of the scale, for example technology, and asked to sort the statements into 11 categories. In the first category should go statements which reflect the most favourable attitude to, for example, technology, in the next category should go items the judge regards as slightly less favourable, the next category should contain slightly less favourable items and so on until category 6, which should contain items which appear neutral. In category 7 should go slightly unfavourable items and so on until category 11, which shows the most unfavourable attitude. Naturally the number of categories could be reduced to 9 or increased to 13, although it is unlikely in the latter case that items could be so finely graded. There have also been modifications to this procedure, such as asking judges to rate items along a 22 cm. line. The intervals between each category should be regarded as equal. The judge is told not to express his own attitude towards the statements, but merely, as objectively as possible, to sort them according to the degree of favourableness of unfavourableness they express. The category a statement has been placed in by a judge should be noted on the

statement, for sorting later.

Once the statements have been ranked by all the judges any set of statements that have not been judged with care should be discarded. So for example, the judgements of any judge who has put an excessive number of statements, say one-third of all statements in any one category should be discarded. Similarly, if a judge appears to have reversed the scales in some instances, putting obviously favourable items in unfavourable categories and vice-versa, then his judgements should be excluded.

If it has been decided to construct several sub-scales towards the attitude object, using the Thurstone method, then judges must be given a set of items for each sub-scale, and must be asked to sort the items into 11 categories for each sub-scale separately. If there is some doubt as to which sub-scale items should be assigned to, before presentation to the judges, a separate group of judges can be used, on an earlier occasion, to sort items into the various sub-scales. (E) (This procedure may be adopted for Likert sub-scales as well.) Gardener suggests that these judges should be given all the items collected, and given a description of the various categories into which they are to sort the items. Once the items have been sorted, a frequency distribution for each item against the possible categories available can be constructed, and items placed in the sub-scale to which the majority of judges have assigned it. Items for which there is a fair degree of disagreement amongst the judges as to which sub-scale they should be assigned to, should be discarded. The scale constructor will have to decide for himself what level of disagreement he will tolerate before discarding an item, based on his own assessment of how many items he requires in each sub-scale item pool, and the overall level of disagreement of the judges.

Another possible way of assigning items to sub-scales would be through factor analysis, but this is a very laborious statistical technique, and almost impossible with 100 items or more. It is in any case unlikely to be more accurate than the method just outlined. However, it may be used to check on the validity of the judges' assignment of items to subscales after the number of items have been reduced by the methods about to be outlined (F). Once this initial judging procedure has been completed, in the

case of Thurstone scales, the item pools for each sub-scale can be presented to a second group of judges, for rating as to the favourability or unfavourability of items within each sub-scale, by the same means as that already described for a single scale.

The procedures which shall now be described for selecting Thurstone items will be with reference to constructing one scale, but the same procedures would have to be adopted for each sub-scale, individually, if several scales were to be constructed.

When constructing a Thurstone scale, what is finally desired is a scale consisting of a limited number of statements, say 20 or less, which sample the whole range of the attitude continuum, which are approximately equally spaced along the attitude continuum, and which are unambiguous. A respondent can only reasonably be expected to respond to a limited number of statements, and therefore those selected should be a representative sample of the possible views a population could have. Statements should be evenly spaced along the attitude continuum, so we can find a person's position on that continuum accurately. Statements should be unambiguous so that when a subject agrees with a statement which, for example, judges have labelled as mildly favourable, we can be reasonably sure that the subject has got a mildly favourable attitude, rather than that he has actually interpreted the statement as very favourable or unfavourable, and therefore actually has an attitude other than mildly favourable.

In order to meet these criteria it is necessary to assign numerical values to the sets of statements returned from the judges. As a statement which the majority of judges have judged to be highly favourable will be found in category 1, the numerical value of that statement on the attitude continuum or scale is considered to be around 1. Similarly, a statement which the majority of judges have judged to be highly unfavourable will be found in category 11, and could be assigned a numerical, or as it tends to be called, scale value of around 11. A neutral statement found in category six would be assigned a scale value of 6. So to give statements a scale value, the scale value of a statement is taken to be the median of the distribution of judgements for each statement i.e. the point on the 11-point scale above and below which 50% of the judgements lie.

In order to select unambiguous statements, statements are selected which the majority of judges place at roughly the same point on the attitude continuum, and can therefore reasonably be assumed to agree as to the favourability or unfavourability of the statement.

Because of this we need to measure the range of categories (out of a possible 11) over which a statement has been placed by the judges, to select those at each position on the attitude continuum which have the smallest range. Any measure of spread could be used for this purpose (e.g. variance, S.D.) but the authors will describe the use of the interquartile range, which is most often used because it is easiest and quickest. The interquartile range is a measure of the middle 50% of judgements. It is called Q and is found by calculating the 25th and 75th centiles. i.e. the position on the attitude continuum below which 25% of judgements lie, and the position on the attitude continuum below which 75% of the judgements lie, and subtracting the value of the former from the value of the latter. When there is good agreement amongst the judges as to the degree of favourability or unfavourability of a statement Q will be small, compared with the value obtained when there is relatively little agreement amongst the judges.

It is then possible to meet the requirements mentioned earlier for a Thurstone scale; we can select a limited number of statements, with scale values covering the whole range of the attitude continuum, scale values approximately equally spaced over the attitude continuum, and which also have small Q values. Thus both scale and Q values are used as criteria for the selection of statements to be included in Thurstone scales. If a choice is to be made among several statements with approximately the same scale values, preference is given to the one with the lowest Q value, i.e. the one believed to be least ambiguous.

Statements can also be selected by finding scale and Q values graphically, using a cumulative proportion graph for each statement. Edwards & Kilpatrick [5] also suggest a method of speeding up the finding of scale and Q values by the graphic method, by taping a master chart of an X and Y axis to a ground glass plate fitted over the top of a wooden box, with a 100 watt bulb inside it.

By placing tracing paper over the chart cumulative proportion graphs for each statement can be quickly drawn and scale and Q values rapidly found from them. If the reader requires the precise details of the mathematical procedures to be undertaken in constructing a Thurstone scale he is advised to consult Edwards,[3] Chapter 4.

5. Further attempts to ensure unidimensionality

To some extent item analysis procedures should have ensured that scales or sub-scales contain, as far as possible, single psychological constructs. Further methods of ensuring the homogeneity of scales are available, however, which would be applicable for Likert as well as Thurstone scales. If authors have constructed several sub-scales and wish to ascertain whether these could be meaningfully aggregated into a single scale, or if authors have constructed a single scale and wish to ensure its homogeneity, one of the following procedures might be adopted. Factor analysis of the items could be undertaken in which case high principal factor concentration would indicate that a single scale would be meaningful, or isolation of separate factors would indicate retention of separate scales. Inter-item correlation matrices or covariance matrices would provide the same information or a coefficient of homogeneity could be utilized such as Loevinger's [11], Scott's [13], or Cronbach's[2] alpha. The recent availability of computers to carry out calculations has made such computation far simpler and the lack of such facilities in the past is probably largely responsible for past scale constructors failing to ensure their scales only measured one psychological construct. For a useful discussion of homogeneity, inter-item correlation and Factor Scoring, see Scott [14].

6. The Final Thurstone Scale

Having selected a small number of statements to form the attitude scale these should be arranged in random order, and presented to subjects with instructions to indicate those they are willing to agree with, and those they disagree with. Taking only the statements with which the respondent has agreed, an

attitude score is obtained from the scale values of these statements, which is regarded as an indication of the location of the subject on the attitude continuum. The attitude score is based on the arithmetic mean or median of the scale values of the statements agreed with.

Two comparable forms of attitude scales are often constructed by those working on Thurstone scales. This is done by selecting from the initial group of statements, in addition to the first set, a second set of statements equal in number to the first, and also having scale values fairly equally spaced along the attitude continuum with fairly low Q values. If both forms of the attitude scale are then given to the same group of subjects the scores for the subjects on the two forms can be correlated, and this correlation taken as a measure of the reliability of the scales.

The final scale or set of sub-scales should be given a pilot run through to check that there are no obvious errors in it, before being tested for reliability and validity. So that, for example, if several respondents agree with 2 or 3 statements with similar scale values, and then one with a widely differing scale value, then the latter statement must be taken out of the scale since it obviously contains elements that are irrelevant to the attitude the scale should be measuring. Even then, the unidimensionality of a Thurstone scale cannot be guaranteed.

As to the reliability and validity of Thurstone scales, the reliability of Thurstone scales tends to be adequate. The avenues available for validation of Thurstone scales are the same as for all attitude scales, and indeed all psychological tests, and the validity of Thurstone scales tends to be no better or worse than that for other attitude scales, or other psychological tests.

Considerable research has been conducted into trying to find out if the attitudes of judges influences the scale values they attribute to statements. Many investigations have indicated that this is not the case, but the work of Hovland and Sherif [9] suggests that this was not revealed by earlier investigations because of the practice of excluding the judgements of judges who place more than 1/3 of statements in one category, apparently through carelessness, which in fact excludes judges with extreme attitudes. This problem could at least be partially overcome if

judges were, where possible, a random sample of the population for whom the final scale is intended, or at least a random sample of adults, rather than, as has often happened, judges being a group of students at the institution at which the scale is being constructed, or some other 'opportunity sample'. At least then it would be likely that favourable, unfavourable, neutral and extreme attitudes would be randomly distributed amongst them, and hence tend to balance out bias in the judgement of statements.

7. The Construction of Likert Scales

Statements collected for Likert Scales for the reasons previously cited should divide into approximately equal classes of favourable and unfavourable statements. It should be noted that the constructor of a Likert scale usually merely assumes that he could obtain agreement to his assignment of statements to favourable or unfavourable categories, and does not seek to discover if others agree with his division. Theoretically it would be possible for other people to see as unfavourable a statement the constructor considers favourable, or vice-versa, but in any case such statements would doubtless be ambiguous and would probably be eliminated by the item analysis procedure.

The entire pool of statements, whether they are to form one scale or several sub-scales, are given to a sample of respondents, for whom the final scale is intended, and the respondents signal their disagreement or agreement by a tick or circle over the appropriate category of "strongly agree", "agree" etc. as already mentioned. At least 100 respondents are generally recommended, though many scale constructors have used 300-400. If the scale is being constructed for young children, and the item pool is large it is probably advisable not to give children the whole item pool at one sitting, but to administer sections of it over a number of occasions, so that it does not become too laborious. After this, if it has been decided to create several sub-scales, then items must be divided up in subscales and item analysis carried out on items within their sub-scales. The method for allocating items to sub-scales is the same as that descibed earlier for Thurstone scales.

The procedures which shall now be described for analysing Likert items will be with reference to constructing one scale,

but the same procedures would have to be adopted for each individual sub-scale, if several were to be constructed.

The scoring method and summation of scores for Likert scales has already been described, so will not be repeated here. In order to decide which are the best statements for a scale, that is which statements best discriminate between those with favourable and those with unfavourable attitudes, and are therefore, also likely to be unambiguous, some form of item analysis is necessary.

To do this scores on the scale under construction are used as a criterion against which to evaluate each statement.

One method of item discrimination, suggested by Edwards [3,4], involves the use of the T-test. We take the 27% of subjects with the highest total score and the 27% of subjects with the lowest total score, and assume that these 2 groups form criterion groups in terms of which to evaluate the individual statements. The mean scores of the high and low groups on each statement are subjected to a T-test. The value of t obtained is a measure of the extent to which a given statement differentiates between the high and low groups. T-values can be looked up in tables of significance, and if the t-value is significant at the 5% level or above this is an indication that the average response of the high and low groups to the statement differs significantly (G). As we finally desire 20-25 highly differentiating statements this can be achieved by picking the 20-25 statements with the largest t-values to form the scale, keeping in mind that roughly ½ the statements should be favourable and ½ the statements unfavourable to the attitude object.

Another method of item discrimination is to use the total score of each respondent as a criterion against which to evaluate their score on each item, by correlating one with the other. The 20-25 statements with the highest overall correlation coefficients can then be retained for the final scale (H). More details of this method can be found in Guildford.[7]

There are other possible methods of item analysis, but Edwards says that it is unlikely that any method of item analysis would result in an ordering of statements essentially different

from any other method of item analysis, so that whichever method appears most convenient should be used. The authors would like to add a point frequently ignored by authors of attitude scales, namely that whatever statistics are used for item analysis, they should be subject to the usual limitations and assumptions of those statistics. So, for example, a t-test should only be used on a sample with normally distributed scores. Authors wishing to establish the homogeneity of their scale or scales can adopt the same procedures here as were described briefly for Thurstone scales under 'Further attempts to ensure unidimensionality.'

8. The final Likert scale

The final set of 20-25 statements is given to a new group of respondents, who are asked to indicate their degree of agreement or disagreement with the statements. An attitude score for each subject is then obtained by summing the weights assigned to the response made to the statements.

As a result of item analysis Likert scales should be unidimensional. However, as Likert scales are not equal-appearing interval scales a subject's score does not represent an absolute position on the psychological continuum. The problem with Likert scales is in interpreting the scores. The same score could be arrived at in a number of different ways and it has therefore been argued that the scores have little meaning. For this reason it has been suggested that the pattern of responses of a respondent should be examined rather than his overall score - no easy task if a large sample is being assessed. It is probably safe to assume that a very low score represents an unfavourable attitude and a very high score represents a favourable attitude, but the interpretation of scores between the maximum and minimum possible scores is difficult. The summated rating score corresponding to neutral is not known, and cannot be assumed to correspond to the midpoint between the high and low scores. The only way to find the neutral point would be to determine the point of least intensity for each scale item, according to the method described by Guttman and Suchmann.[8]

It would also be possible for a respondent to agree with half of the favourable and unfavourable items, and disagree with

the other half of the favourable and unfavourable items, thereby achieving a score representing an approximately neutral attitude, when his attitude is anything but neutral. By examining the pattern of responses of subjects, papers with inconsistent responses can be observed and such papers discarded. However, the authors do wonder if this action is justified, as it seems that the majority of people have ambivalent feelings towards most controversial issues, and so inconsistent responses may be very consistent with a person's feelings.

Apart from the difficulty of inconsistent responses, as long as our purpose is not to assign a single person to a class of those favourable or unfavourable in attitude, (when a Thurstone scale would be more appropriate) uncertainty about interpreting the middle range of summated rating scores will not present a problem. If our interest is in comparing the mean change in attitude scores as a result of introducing some experimental variable, or comparing the mean attitude scores of 2 or more groups, or correlating scores on an attitude scale with other personality measures, then knowledge of a midpoint is unnecessary. Likert scales are thought to do very well when it comes to reliable, rough ordering of people with regard to a particular attitude, and Likert scales do provide two advantages Thurstone scales don't have. They provide more precise information about a respondent's degree of agreement or disagreement with scale items, and respondents usually gain more satisfaction from a 5-point degree of agreement or disagreement answering mode, rather than giving a simple agree/disagree response. In any case, if we do wish to compare the attitude of a single subject relative to the attitude of other subjects, then percentage or standard deviation norms can be calculated, if a sample of sufficient size is available.

From the evidence, reliability of Likert scales, tested by the split-half or test-retest method, is good and often better than that of Thurstone scales. (I). As Likert scales are not equal-appearing interval scales Likert-selected items do not fall at equally spaced intervals along the attitude continuum. However, this does not necessarily show that the prime purpose of attitude scales, the ordering of individuals with respect to the degree of favourability or unfavourability they associate with

a psychological object, is not being met. Edwards and Kenney,[4] compared the ordering of individuals by a Thurstone and a Likert scale, independently constructed from the same original set of statements, using a number of university Psychology students. The correlation coefficient between the scores on the Likert scale and the Thurstone scale was found to be extremely high. This evidence suggests that the relative ordering of subjects on either equal-appearing interval scales or summated rating scales is, for all practical purposes, essentially the same.

9. Comparison of ease with which Likert and Thurstone scales are constructed

Likert claimed that his method of constructing an attitude scale was quicker and simpler than Thurstone's. Edwards and Kenney[4] estimate that constructing an equal-appearing interval scale takes about twice as long as that of constructing a summated rating scale. However, there are various methods by which the Thurstone procedure can be speeded up, and these will be briefly described.

Seashore and Hevner[15] used a judgement procedure by which judges rated each statement on a 9-point scale, A, B, C, D, E, F, G, H, I, which was printed on the left hand side of each item. They estimated that the procedure saved from 50% to 87% of time on the various processes involved. They reported that the subjects also found the task more pleasant than Thurstone's original procedure, and when the results of a group of subjects who had rated statements by their method, and a group of subjects who had rated the statements by Thurstone's method were compared, only negligible differences in the medians, scale, and Q values of items were found.

Ballin and Farnsworth[1] used a graphic rating method for obtaining judgements, using statements taken from an original Thurstone scale. Judges were asked to rate statements along a 22 inch line. Comparison of score values involved measuring the distance between the unfavourable end and the point at which the given statement was placed by the rater. These values were multiplied by a constant to make them comparable with Thurstone

scale values. A separate group of judges was asked to rate the statements by the Seashore and Hevner method. Surprisingly close agreement was found between the scale values obtained by the method of graphic rating, the Seashore and Hevner method, and the original Thurstone equal-appearing interval method. Ballin and Farnsworth suggest that a 22 cm. line might prove more useful than a 22 inch line, because measurement in centimetres can simply be divided by 2 to obtain values comparable to the Thurstone method, and the shorter line might improve judgements by enabling the rater to perceive more of the attitude continuum in a single eye-span.

Once judgements have been obtained, scale and Q values can be rapidly determined through the use of a nomograph such as Jurgensen's [10], or by a master chart such as Edwards and Kilpatrick [5] used, mentioned earlier. It seems, therefore, that by using such aids the method of summated ratings and the method of equal-appearing intervals would be fairly comparable with respect to time and labour involved. With the use of computers, of course, the calculations involved in constructing either type of scale are drastically reduced and consequently the decision as to what type of scale to construct should no longer rest on considerations of time. Instead, having described the construction of Thurstone and Likert scales, it is probably more apparent to the reader that the scale which a researcher would use would depend on the purpose of the scale. Scales have been developed to concentrate on fulfilling certain requirements of an attitude scale, whilst paying less attention to other requirements. Depending on the use made of a scale will depend which type of scale is most appropriate. Confining ourselves to Thurstone and Likert scales, if we wish to study the pattern of an attitude, which sub-areas of an attitude are dominant, and which areas are unimportant, the Likert procedure may be best utilized. If we are to study the difference in attitudes between people then Thurstone procedures may prove better. If we wish to divide people broadly into 2 or 3 groups with favourable, unfavourable and neutral attitudes, the Likert procedure would be adequate, but if we wished to divide groups more accurately into degrees of favourability or unfavourability then the Thurstone procedure might be more appropriate. Other methods might suit other purposes. The authors chose to

construct an attitude scale using the Likert method because the scale was only to be used to divide pupils broadly into 2 groups, and also because it was felt that children would respond better to having a 5-point choice of response. (J).

Similarly, the way in which a scale is validated may depend upon its purpose. If it is to predict something, then predictive validity would need to be proved, if it was to divide people into groups then construct validity would have to be determined. (K) (L).

A P P E N D I X 1

The development of the technology scale

THE TECHNOLOGY QUESTIONNAIRE - A BRIEF SUMMARY OF ITS DEVELOPMENT

Statements of attitude towards technology were collected from over 400 secondary school children and sorted by judges into four Likert-type attitude sub-scales. The most discriminating items from each scale were chosen for the four final scales on the basis of the responses of a second sample of 101 fourth year boys. Satisfactory reliability and homogeneity was indicated for the four scales, and factor analysis confirmed the isolation of four separate constructs requiring separate scales. Scores on all scales were found to correlate positively and significantly with the number of science and technical options chosen by school pupils. This provides some validity for the four scales since attitude towards technology is likely to influence the subjects school children choose to study. It is concluded that the scales form a reliable andvalid instrument which is presently being used to assess attitudes toward technology.

Introduction

There has been much concern in recent years regarding the low level of industrial awareness of children leaving our schools and more particularly their seemingly negative attitudes to the role of technology and industry in our society. For example, the criticism has been made that schools have actively formed and reinforced negative attitudes in the minds of pupils towards such things as technological development and industrial careers. It would seem that many young people do not see technology as a poverty reducing agent but rather as a necessary evil and reject a career in industry because of the low esteem in which it is held. Many suggestions have been put forward to account for this state of affairs, but at best many of these are purely speculative, apart from one or two isolated reports, e.g. Page[17], Hutchings[18] and Dainton[19] This is undoubtedly due to the paucity of test instruments in this field and with this in mind, the authors decided in 1976 to develop an attiude test which could assess specific attitudes towards technology and industry. It was felt that the greatest need for assessment of such attitudes lay in the 13-18 age range and so the resulting test covers this age band.

NATURE AND TYPE OF SCALE

Since the scale under discussion was to be used from the point of view of placing 13 to 18 year old students into a number of broad groups with respect to the four facets of attitude towards technology which it purports to assess, the authors decided that a likert-type scale was most suitable.(J). The scale also offers respondents five categories of response from strongly agree to strongly disagree, and does not rely on extensive and lengthy judging procedures, but rather uses the respondents' own endorsements as a basis for item evaluation.

PREPARATORY INTERVIEWS OF TARGET POPULATION SAMPLES

(A)The purpose of the initial interviews was to ascertain whether pupils in the age range 13-18 years:

(a) used the word 'technology' and had some understanding of its meaning;
(b) used the word 'industry' and had some understanding of its meaning;
(c) could converse freely about technology, the work of technologists, technical training and education at tertiary level, and industrial careers.

In all the above, pupils demonstrated their ability to discuss such matters and it was noticeable how many strong and unsolicited opinions emerged during the course of the interviews. It was therefore decided to conduct longer interviews with pupils drawn from broad samples to facilitate the collection of a pool of statements which could be sorted and considered for inclusion in the projected attitude measuring instrument.

(B) Over 400 pupils, (boys and girls), were interviewed. They ranged from 13 to 18 years and covered the full ability range. Subjects were drawn from two large mixed comprehensive schools, one receiving its pupils from an urban catchment area, the other drawing a substantial number of its pupils from a rural area. The major proportion of these interviews was conducted in the urban school. Extra subjects were interviewed in the rural school to see whether the statements collected from these pupils differed significantly from those previously collected. In fact this was

not the case and the item pool collected from the second school only reproduced that collected from the first school.

(B) The resulting interview recordings were edited and statements thought to be relevant to the attitude areas under study were collected. These items were sorted and categorized. Four major categories were identified and items within these sets were ascribed a face validity for their category:

(a) Those deemed to be representative of an attitude to Technology and Technologists.

(b) Those deemed to be representative of an attitude to careers in Industry.

(c) Those deemed to be representative of an attitude to Technical Training and Education at tertiary level.

(d) Those deemed to be representative of an attitude to Technology as a new school subject.

All other statements were discarded.

(D) The subsequent item pools for all four constructs were then subjected to scrutiny, using Edwards [3] criteria for attitude statements. All factual, ambiguous, non-discriminatory or otherwise unsuitable statements were eliminated. The remaining 106 items were selected for possible inclusion in the trial instrument.

ITEM ANALYSIS

(E) Whilst the trial instrument was being tested, the four construct subscales were described to ten judges, (five teachers, five fifth-year pupils). They were required to judge the items, which had been supplied to them on separate slips of paper, and to ascribe them to any one of the constructs as defined above, or to a fifth category, where the item was deemed to be either ambiguous, or was judged to be of another attitude trait.

Eleven items failed to gain unanimous voting and were therefore judged to be unsuitable for inclusion in the final scales. Also, they were not included in the statistical item analysis that was to follow.

One hundred and one fourth-year boys of known interests, and considered to be a fair sample of the target population, were tested with the trial instrument and their responses formed the basis for the detailed statistical item analysis.

(H) The original item pools were purposely made large, so that weak items could be eliminated and the ten best items could go forward for inclusion. As an ideal, each item should correlate with a reliable outside criterion of the attitude that it is supposed to measure. Such criteria of high reliability are difficult to find with regard to attitude measures and therefore the best available measure is usually considered to be the item pool. To this end, item total correlations on each subscale were obtained and the 10 items with the highest correlations within each subscale were chosen for inclusion in the final scale. (G). To ensure that these items discriminated sufficiently well, a t-test was carried out on the mean scores on each item of the highest scoring 27% and the lowest scoring 27% for the subscale in question. This was an additional check as items with the highest item-total correlations usually have associated high discrimination indexes.
As a result of this selection procedure, the four subscales were found to be homogeneous and free from weak discriminatory items. The final 40 items comprising the 4 scales were randomly ordered and printed in a booklet (see Appendix 2).

SCALE RELIABILITY (I)

Table 1 gives the Cronbach[2] alpha coefficients that were gained for the four reduced subscales. Taking Cronbach's coefficient to represent the mean of all possible split-half correlation coefficients, the values listed in Table 1 suggest that each subscale has a satisfactory reliability and homogeneity.

TABLE I

Subscale	Cronbach's Alpha Coefficient
Attitude to Technology	.84
Attitude to an Industrial Career	.83
Attitude to Technical Training	.82
Attitude to Technology as a School Subject	.38

CONSTRUCT VALIDITY

The internal consistency of judged item sets was seen as a measure of construct validity. However, as a further check on the item sets, and also on the relative uniqueness of the four constructs, items were factor analysed. As some degrees of association between the scales might be expected, an Oblique Solution with iterations was applied and rotated. The results are set out in Table ?. Loadings less than 0.35 were omitted and five main factors were interpreted.

As expected, some degree of overlap was evident between the scales but, interestingly, not in the case of the Attitude to a Career in Industry. This scale is, as later work was to demonstrate, virtually independent of the other three

Of the ten items from the Attitude to Technology scale, nine loaded highly and positively on factor IV, although a few loadings, generally of lesser values, cross-referenced on factors 1 and 111. On the Attitude to Careers in Industry Scale, the ten items loaded substantially on factors 11 and 1V1, suggesting by this structure that two factors are at work in this scale. The least distinct structure was produced on the Attitude to Technical Training Scale, although loading was, if anything, biased most heavily on factor 111. The degree of cross-referencing on this scale would suggest that a general technology factor is operating here, in that it would seem to be picking up an overall attitude. It is interesting to note that this scale overlaps on all other scales, except Attitude to a Career in Industry. On the Attitude to Technology as a School Subject Scale, all ten items loaded substantially on factor 1. Cross-references on factors 111 and 1V were small in number and generally of lower values.

As a general comment on these results, it appears that the factor analysis has in the main given support to the judged item sets. Only in the case of the Career in Industry Scale is a different structure evident. Since this scale is clearly unique, and since four of the items are common to both factors, separation was not considered viable. It was therefore felt that the scale should be held together as judged for the purpose of

analysis. The factor structure has shown that the Attitude to Technical Training is the weakest of four scales. This had been indicated previously, by the slightly lower alpha obtained on this scale as compared with the remaining three.

Table 2

FACTOR ANALYSIS OF ATTITUDE ITEMS TO CONFIRM JUDGED ITEM SETS

SCALE	ITEM	I	II	III	IV	V	VI	VII
ATTITUDE TO TECHNOLOGY	4				44			
	5				52			
	6							
	11				45			
	18	43			61			
	21				52			
	28			39	62			
	32	37			59			
	38	42			58			
	40			47	36			
ATTITUDE TO CAREERS IN INDUSTRY	2					54		
	3					52		
	8					65		
	22		37					
	24		58			41		
	27		57			41		
	31		43			46		
	34		75					
	35		74					
	36		64			39		
ATTITUDE TO TECHNICAL TRAINING	1							
	13	41		43	37			
	15	58		43	42			
	17	39		50				
	19	40		50				
	20	38			43			
	23			48				
	29				46			35
	30			60				
	39			69				
ATTITUDE TO TECHNOLOGY AS A SCHOOL SUBJECT	7	68						
	9	67			40			
	10	42						
	12	67						
	14	67		38				
	16	65						
	25	57		36				
	26	44			43			
	33	63			49			
	37	58						

OBLIQUE. Seven factors extracted and rotated. Decimal points omitted. Loadings less than $\pm$0.35 omitted.

Further evidence of validity was sought for the four subscales by establishing two criteria groups for the 101 respondents on the following basis :

Group A : Pupils who had freely opted to pursue three or more science and technical options, (not including biology), at 4th and 5th year level.

Group B : Pupils who had opted to take one or less science and technical courses at this level, (not including biology).

True to expectation, analysis of the four subscales showed that the mean scores gained by Group A were significantly greater than those obtained by Group B, as Table 3 shows.

TABLE 3

BISERIAL CORRELELATION COEFFICIENT BETWEEN MEAN SCORES OBTAINED BY CRITERIA GROUPS A AND B USING THE REDUCED (FINAL) SUBSCALES

Attitude Scale	Biserial Correlation Coefficient	
Attitude to Technology	.45p	.01
Attitude to Career in Industry	.49p	.001
Attitude to Technical Training	.42p	.01
Attitude to Technology as a School Subject	.62p	.001

(K) As a finer check, criteria ratings ranging from 0 for those not taking physics, chemistry, integrated science or technical subjects as 4th and 5th year options to a maximum of 4 were awarded to respondents. It was found that the higher the criteria rating pupils had (ie. the more technologically oriented subjects they were studying) the higher the subscale scores, as the correlations of Table 4 indicate.

TABLE 4

CORRELATIONS BETWEEN ATTITUDE SUBSCALES AND CRITERIA RATING IN ORDER OF MAGNITUDE

Pairs	Pearson's Correlation Coefficient
Criteria Rating with Attitude to Technology as a School Subject	.41 p .001
Criteria Rating with Attitude to Technology	.35 p .001
Criteria Rating with Attitude to Career in Industry	.30 p .01
Criteria Rating with Attitude to Technical Training	.25 p .01

It seems reasonable to suppose that the attitude of pupils towards facets of technology is likely to influence the amount of technology they choose to have in their curriculum when subject options are available.

It is possible to present the results of Table 4 in a slightly different way so that the mean scores of respondents in each criteria group are presented so that comparisons can be made with other samples. Table 5 gives this information. It should be

noted that the small number of respondents in group 4 make these averages slightly suspect.

TABLE 5

MEAN SUBSCALE SCORES FOR THE CRITERIA GROUPS OF TABLE 4

Criteria Group	Attitude to Technology	Attitude to Industrial Career	Attitude to Technical Training	Attitude to Technology as a School Subject
0 & 1 n=38	25.47	23.97	25.76	22.79
2 n=35	29.97	26.37	28.84	28.71
3 n=20	30.05	30.07	29.45	31.85
4 n=9	35.40	33.00	34.00	34.00

(L) A group of 44 British Rail apprentices has been tested against Table 5 and their results (shown in Table 6) indicate that their attitude levels compare favourably with the respondents of Table 5 and suggests candidates of suitable calibre are being recruited. This provides further evidence of the construct validity of the four sub-scales.

TABLE 6

BRITISH RAIL APPRENTICE ATTITUDE SCORES

Sample Size	Attitude to Technology	Attitude to Industrial Career	Attitude to Technical Training	Attitude to Technology as a School Subject
n = 44	29.45	30.30	30.00	29.16
Criteria Group	2	3	3	2

The slightly stronger attitude to a career in industry and technical training is to be expected, but the less strong attitude to technology may be indicative of future career dissatisfaction and uncertainty.

All four subscales were also to correlate significantly with one another, as can be seen in Table 7.

TABLE 7

CORRELATIONS OF SUBSCALES IN ORDER OF MAGNITUDE

Subscale Pairs N = 101	Pearson's Correlation Coefficient
Attitude to Technical Training with Attitude to Technology as a School Subject	.64 p .001
Attitude to Technology with Attitude to Career in Industry	.60 p .001
Attitude to Career in Industry with Attitude to Technology as a School Subject	.54 p .001
Attitude to Technology with Attitude to Technical Training	.54 p .001
Attitude to Technology with Attitude to Technology as a School Subject	.48 p .001
Attitude to Career in Industry with Attitude to Technical Training	.48 p .001

CONCLUSION AND DEVELOPMENTS

An instrument, comprising the four subscales as described above, has been developed, tried and validated, using a fair sample of the target population of 13-18 year olds for whom it is intended. It has been established that the instrument is well within the capacity of the majority of fourth-year pupils and can be completed well within the projected 30 minutes time limit imposed. It has been developed to yield four distinct attitude measures and it has been established that such measures can be interpreted and used as a basis for group comparisons. Already the attitude scale has been used to compare railway apprentices scores with the criteria means of table 5. It is also being used in the evaluation programme of the Avon, Bromley, Bath University, Schools Council Modular Courses in Technology. Further, a project funded by the Department of Industry is working to establish regional norms for the scales and also to identify schools with high positive scores for special study. To date some 14,000 children have been tested using the scales and the results will be available sometime in 1979-80.

A P P E N D I X 2

Technology Attitude Scale

TECHNOLOGY QUESTIONNAIRE

This questionnaire forms part of a national survey to find out how secondary school pupils feel about technology. You are not asked to give your name so any information you provide will be confidential.

In the last century, many changes have come about by the development and use of TECHNOLOGY. Many people, such as fuel and rubber technologists, engineers, designers, metallurgists, architects, townplanners and others have used their knowledge of the world about them to solve the problems they see. All these people can be described as TECHNOLOGISTS.

Although, compared with our great-grandparents' days, we can see many improvements in our modern way of life, such things as the threat of nuclear war and the pollution of our environment are also the results of TECHNOLOGY.

We have made a collection of the things which people have said about TECHNOLOGY AND TECHNOLOGISTS and we should like to know which statements you agree with.

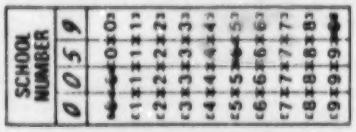
This is not a test and the way you answer the following statements cannot affect your results on school courses. Neither are there any right or wrong ways of answering. All we want you to do is to go through the statements in the questionnaire and mark each one according to the way you feel about it.

UNIVERSITY OF BATH, SCHOOL OF EDUCATION

RP © M077
MN

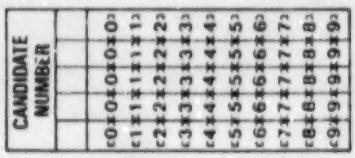
Please fill in the following details about yourself, by pencilling in the appropriate box(es). If, for example, your school number is 0059 you would fill in the boxes thus:

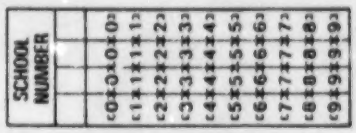
EXAMPLE

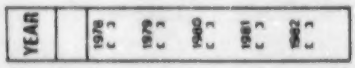


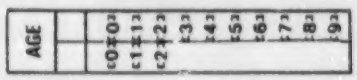
Please ask your teacher if you are unsure how to reply. If you make a mistake rub it out with a rubber and enter new answer.

USE AN HB PENCIL









Please pencil in the box besides each subject you are presently studying at school. For subjects not studied leave a blank box.

GENERAL, OR SECONDARY, OR INTEGRATED SCIENCE		MOTOR ENGINEERING	
MATHS		TECHNICAL DRAWING (OR ENGINEERING DRAWING)	
PHYSICS AS A SEPARATE SUBJECT		TECHNICAL DESIGN (NOT ART)	
CHEMISTRY AS A SEPARATE SUBJECT		TECHNOLOGY	
BIOLOGY AS A SEPARATE SUBJECT		WOODWORK	
ENGINEERING SCIENCE		METALWORK	
ENGINEERING (NOT METALWORK)		APPLIED SCIENCE	

	Yes	No	Don't Know
WOULD YOU LIKE A TECHNOLOGICAL CAREER?			

QUESTIONNAIRE

Please show your agreement or disagreement with each of the statements by pencilling in the box underneath the most suitable letter, according to the code below, USING AN HB PENCIL.

- 'A' – I strongly agree with this statement
- 'a' – I almost agree, but not completely
- '?' – I am undecided, or unsure
- 'd' – I mainly disagree, but not completely
- 'D' – I disagree strongly

FOR EXAMPLE

	A	a	?	d	D
'Doughnuts are horrible'					■

I strongly disagree with this statement, therefore I have pencilled under the letter 'D'.

Note that the code letter sequence is not the same for all questions.

1.	It is important to our country that as many of us as possible get a good Technical training.	D	d	?	a	A	4.	Technologists cause more problems than they solve	A	a	?	d	D
2.	People who work in Industry have to work in stuffy, smelling places.	A	a	?	d	D	5.	I think if we all tried to live without Technology, we would be much happier as people.	A	a	?	d	D
3.	I do not see how anyone could feel happy working in Industry.	A	a	?	d	D							

Continued on next sheet

Precision printed by Reedform Ltd. Windsor, England

REFERENCES

REFERENCES

1. Ballin, M. and Farnsworth, P.R. 'A graphic rating method for determining the scale values of statements in measuring social attitudes.' Journal of Social Psychology, 1941, 13, pp. 323-327.

2. Cronbach, L.J. 'Coefficient alpha and the internal structure of tests.' Psychometrika, 1951, 16, 3, pp.297-334.

3. Edwards, A.L. 'Techniques of attitude scale construction.' New York: Appleton Century 1957.

4. Edwards, A.L. and Kenney, K.C. 'A comparison of the Thurstone and Likert techniques of attitude scale construction.' Journal of Applied Psychology, 1946, 30, pp. 72-83.

5. Edwards, A.L. and Kilpatrick, F.P. 'A technique for the construction of attitude scales. Journal of Applied Psychology, 1948, 32, pp. 374-384.

6. Gardner, P.L. 'Attitude Measurement'. Educational Research, 1975, 17, 2.

7. Guilford, J.P. 'Fundamental Statistics in Psychology and Education.' McGraw-Hill 1965.

8. Guttman, L. and Suchman, E.A. 'Intensity and a zero point for attitude analysis.' American Sociological Review, 1947, 12, pp. 57-67.

9. Hovland, C.I. and Sherif, M. 'Judgement phenomena and scales of attitude measurement: item displacement in Thurstone scales.' Journal of Abnormal Psychology, 1952, 47, pp.822-832.

10. Jurgensen, C.E. 'A nomograph for rapid determination of medians'. Psychometrika, 1943, 8, pp. 265-269.

11. Loevinger, J. 'The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis'. Psychological Bulletin, 1948, 45, pp. 507-529.

12. Oppenheim, A.N. 'Questionnaire design and attitude measurement' Heinemann, 1966.

13. Scott, W.A. 'Measures of test homogeneity' Educational and Psychological Measurement, 1960, 20, 4, pp. 751-757.

14. Scott, W.A. 'Attitude Measurement' in Lindzey, G. and Aronson, E. (eds). The Handbook of Social Psychology, Vol 2 (2nd Edition). Addison Wesley, 1968.

15. Seashore, R.H. and Hevner, K. 'A time-saving device for the construction of attitude scales.' Journal of Psychology,1933, 4, pp. 366-372.

16. Shaw, M.E. and Wright, J.M.'Scales for the Measurement of Attitudes.' McGraw Hill 1967.

17. Page, G.T. 'Engineering Among the Schools' London; Institute of Mechanical Engineers, 1965.

18. Hutchings, D.J. 'Technology and the 6th Form Boy' Oxford University, Institute of Education, 1963.

19. Dainton Report 'Enquiry into the Flow of Candidates in Science and Technology into Higher Education' H.M.S.O., CMND 2893, 1966.

CANDIDATE NUMBER

SCHOOL NUMBER

QUESTIONNAIRE (continued)

6. It is a good thing we put restrictions on some of the things Technologists do, because otherwise they would just wreck everything that really matters in life.	A	a	?	d	D
7. I would hate to have to learn about Technology as a school subject.	A	a	?	d	D
8. Work in Industry is dirty and unpleasant.	A	a	?	d	D
9. Technology would be a good subject to study in school.	D	d	?	a	A
10. Technical subjects at school are just too hard to understand.	A	a	?	d	D
11. Our Technologists are very clever and responsible people.	D	d	?	a	A
12. Technology as a school subject would be of no benefit to me.	A	a	?	d	D
13. A Technical training makes you a very useful and practical person.	D	d	?	a	A
14. I would love to do more Technical courses at school.	D	d	?	a	A
15. A good Technical training is well worthwhile.	D	d	?	a	A
16. Technical subjects at school are all right, but I would not choose to do any of them.	A	a	?	d	D
17. Getting a Technical training is one of the really good things a person can do after leaving school.	D	d	?	a	A
18. I think people would be far better off without all this Technology.	A	a	?	d	D
19. A Technical training helps you to do lots of really exciting things.	D	d	?	a	A
20. With the job situation as it is, only a fool would spend his time on a Technical training.	A	a	?	d	D
21. Technology does not serve us as it should, but makes us live in ways we don't like.	A	a	?	d	D
22. To work in Industry is a good way of using your education for the benefit of the country.	D	d	?	a	A
23. Having a Technical training means that you can get a job almost anywhere.	D	d	?	a	A
24. Working in Industry is monotonous and boring.	A	a	?	d	D
25. I would really like to learn more in school about the way in which our Technologists solve problems.	D	d	?	a	A
26. The average pupil would gain nothing by having to learn Technology.	A	a	?	d	D
27. Working in Industry all your life would be like one big nightmare.	A	a	?	d	D
28. Technologists are very useful people.	D	d	?	a	A
29. It is stupid to get a Technical training, because it does not enable you to get good jobs.	A	a	?	d	D
30. People who have had a Technical training usually do well in life.	D	d	?	a	A
31. People who work in Industry become unhealthy because of their work.	A	a	?	d	D
32. Technology is unnecessary and wastes our resources.	A	a	?	d	D
33. I can see no value in learning about Technology at school.	A	a	?	d	D
34. If I end up in Industry, it will be a waste of all my education.	A	a	?	d	D
35. Working in Industry has nothing to offer me.	A	a	?	d	D
36. You are too shut in when you work for Industry.	A	a	?	d	D
37. Learning about Technology is not important to my education.	A	a	?	d	D
38. I don't agree with all this Technology. We should go back to a simpler way of life.	A	a	?	d	D
39. I think people should go for a technical training because in the end they will get the best jobs.	D	d	?	a	A
40. If we had more Technologists it would greatly improve our way of life.	D	d	?	a	A

PUBLICATIONS AVAILABLE FROM SCIENCE & TECHNOLOGY EDUCATION CENTRE

Strategies for Change in a Technological Society	£1.00	
Attitude Assessment in Science & Technology	£2.00	
An Evaluation of the Use by Teachers of British Petroleum's Multi-media Resource Pack: "A North Sea Adventure"	£2.00	
The Effects of Microprocessors on Industry, Society and Employment	£1.00	
Schools Council Modular Courses in Technology		
Progress Report No. 1	Free	*
Progress Report No. 2	Free	*
A Brief Look at Mathematics in the Engineering Industry	Free	*
A Small Survey of the Use of Mathematics in The Engineering Industry	Free	*
A Survey of Mathematics Projects involving Education and Employment	Free	*
An Evaluation of the Use of Self-instructional Material in the Library	£3.00	
Bulletin (published three times each year)	Free	*
Assessment in Higher Education (an international journal published three times each year)	£6.50	

All prices are U.K. prices

* Limited Supplies

For further details write to:

Science & Technology Education Centre
University of Bath
Bath
BA2 7AY
U.K.